Chapter 4

Continuous Variables and Their Probability Distributions

4.2 - The probability distribution for a Continuous Random Variable

- Cumulative distribution Function

Let Y denote any random variable. The distribution function of Y, denoted by $F(y)$, is such that

$$F(y) = P(Y \leq y) \text{ for } -\infty < y < \infty$$

• The nature of the distribution function associated with a random variable determines whether the variable is continuous or discrete.

EXAMPLE 4.1 Suppose that Y has a binomial distribution with $n = 2$ and $p = 1/2$. Find $F(y)$.

Solution The probability function for Y is given by

$$p(y) = \binom{2}{y}\left(\frac{1}{2}\right)^{y}\left(\frac{1}{2}\right)^{2-y}, \qquad y = 0, 1, 2,$$

which yields

$$p(0) = 1/4, \qquad p(1) = 1/2, \qquad p(2) = 1/4.$$

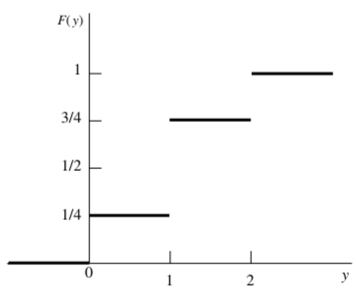


FIGURE **4.1** Binomial distribution function, $n = 2$, $p = 1/2$

What is $F(-2) = P(Y \le -2)$? Because the only values of Y that are assigned positive probabilities are 0, 1, and 2 and none of these values are less than or equal to -2, $F(-2) = 0$. Using similar logic, $F(y) = 0$ for all $y < 0$. What is $F(1.5)$? The only values of Y that are less than or equal to 1.5 and have nonzero probabilities are the values 0 and 1. Therefore,

$$\begin{aligned} F(1.5) = P(Y \le 1.5) &= P(Y = 0) + P(Y = 1) \\ &= (1/4) + (1/2) = 3/4. \end{aligned}$$

In general,

$$F(y) = P(Y \le y) = \begin{cases} 0, & \text{for } y < 0, \\ 1/4, & \text{for } 0 \le y < 1, \\ 3/4, & \text{for } 1 \le y < 2, \\ 1, & \text{for } y \ge 2. \end{cases}$$

A graph of $F(y)$ is given in Figure 4.1. ■

Properties of a Distribution Function

If $F(y)$ is a distribution function, then:

1. $F(-\infty) = \lim_{y \to -\infty} F(y) = 0$

2. $F(\infty) = \lim_{y \to \infty} F(y) = 1$

3. $F(y)$ is a non-decreasing function of y.
[If y_1 & y_2 are any values such that $y_1 < y_2$ then $F(y_1) \leq F(y_2)$].

DEFINITION **4.2** A random variable Y with distribution function $F(y)$ is said to be *continuous* if $F(y)$ is continuous, for $-\infty < y < \infty$.[2]

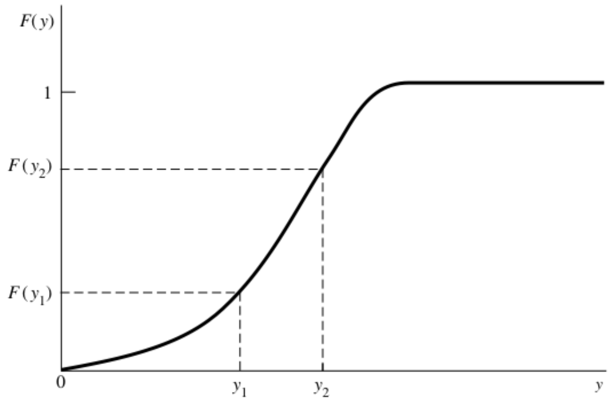


FIGURE **4.2** Distribution function for a continuous random variable

Note:

To be mathematically precise, we also need the first derivative of $F(y)$ to exist and to be continuous, except, for, at most a finite number of points in any finite interval.

If Y is a continuous random variable then for any real number y,

$$P(Y=y)=0.$$

$\cup$

→ If this were <u>not</u> true and $P(Y=y_0)=p_0>0$,

then $F(y)$ would have a discontinuity (jump) of size p_0 at the point y_0, violating the assumption that Y was continuous.

- The derivative of $F(y)$ is another function of prime importance in probability theory & statistics.

DEFINITION 4.3

Let $F(y)$ be the distribution function for a continuous random variable Y. Then $f(y)$, given by

$$f(y)=\frac{dF(y)}{dy}=F'(y)$$

wherever the derivative exists, is called the *probability density function* for the random variable Y.

↳ It follows from Definitions 4.2 & 4.3

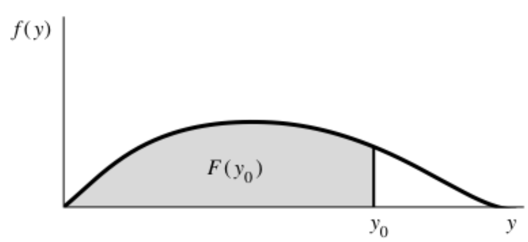
that $F(y)$ can be written as:

$$\boxed{F(y) = \int_{-\infty}^{y} f(t)\,dt.}$$

where: $f(\circ)$: probability density function
t : variable of integration.

Graphical relationship between the distribution & density function

FIGURE 4.3 The distribution function



* The probability density function is a **theoretical model**, for the frequency distribution (histogram) of a

population of measurement.

Properties of a Density Function

- If $f(y)$ is a density function for a continuous random variable, then

1. $f(y) \geq 0$ for all y, $-\infty < y < \infty$

2. $\int_{-\infty}^{\infty} f(y)\,dy = 1$

→ Because $F(y)$ is a non-decreasing function, the derivative $f(y)$ is **never negative**.

→ $F(\infty) = 1$ & therefore, $\int^{\infty} f(t)\,dt = 1$.

EXAMPLE 4.2 Suppose that

$$F(y) = \begin{cases} 0, & \text{for } y < 0, \\ y, & \text{for } 0 \le y \le 1, \\ 1, & \text{for } y > 1. \end{cases}$$

Find the probability density function for Y and graph it.

Solution Because the density function $f(y)$ is the derivative of the distribution function $F(y)$, when the derivative exists,

$$f(y) = \frac{dF(y)}{dy} = \begin{cases} \dfrac{d(0)}{dy} = 0, & \text{for } y < 0, \\ \dfrac{d(y)}{dy} = 1, & \text{for } 0 < y < 1, \\ \dfrac{d(1)}{dy} = 0, & \text{for } y > 1, \end{cases}$$

and $f(y)$ is undefined at $y = 0$ and $y = 1$. A graph of $F(y)$ is shown in Figure 4.4.

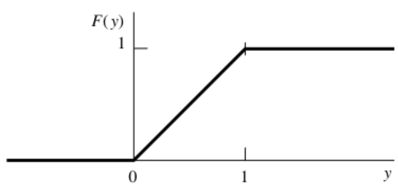


FIGURE **4.4** Distribution function $F(y)$ for Example 4.2

■

EXAMPLE 4.3 Let Y be a continuous random variable with probability density function given by

$$f(y) = \begin{cases} 3y^2, & 0 \le y \le 1, \\ 0, & \text{elsewhere.} \end{cases}$$

Find $F(y)$. Graph both $f(y)$ and $F(y)$.

Solution The graph of $f(y)$ appears in Figure 4.6. Because

$$F(y) = \int_{-\infty}^{y} f(t)\,dt,$$

we have, for this example,

$$F(y) = \begin{cases} \int_{-\infty}^{y} 0\,dt = 0, & \text{for } y < 0, \\ \int_{-\infty}^{0} 0\,dt + \int_{0}^{y} 3t^2\,dt = 0 + t^3\big]_0^y = y^3, & \text{for } 0 \le y \le 1, \\ \int_{-\infty}^{0} 0\,dt + \int_{0}^{1} 3t^2\,dt + \int_{1}^{y} 0\,dt = 0 + t^3\big]_0^1 + 0 = 1, & \text{for } 1 < y. \end{cases}$$

Notice that some of the integrals that we evaluated yield a value of 0. These are included for completeness in this initial example. In future calculations, we will not explicitly display any integral that has value 0. The graph of $F(y)$ is given in Figure 4.7.

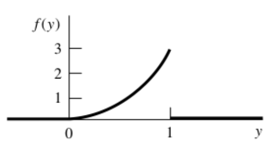


FIGURE **4.6** Density function for Example 4.3

■

Definition 4.4

- Let Y denote any random variable. If $0 < p < 1$, the **p^{th} quantile of Y** denoted by ϕ_p

→ is the smallest value that
$$P(Y \leq \phi_q) = F(\phi_p) \geq p.$$

If Y is continuous, ϕ_p, is the smallest value such that
$$F(\phi_p) = P(Y \leq \phi_p) = p$$

[Some prefer to call ϕ_p the $100p^{th}$ percentile of Y].

* A important special case *

When $p = 1/2$, and $\phi_{0.5}$

median

* $P(a < Y \le b) = P(Y \le b) - P(Y \le a)$

$$= F(b) - F(a)$$

$$= \int_a^b f(y)\, dy$$

THEOREM **4.3** If the random variable Y has density function $f(y)$ and $a < b$, then the probability that Y falls in the interval $[a, b]$ is

$$P(a \le Y \le b) = \int_a^b f(y)\, dy.$$

If Y is a continuous random variable and a and b are constants such that $a < b$, then $P(Y = a) = 0$ and $P(Y = b) = 0$ and Theorem 4.3 implies that

$$P(a < Y < b) = P(a \le Y < b) = P(a < Y \le b)$$
$$= P(a \le Y \le b) = \int_a^b f(y)\, dy.$$

The fact that the above string of equalities is *not*, in general, true for discrete random variables is illustrated in Exercise 4.7.

4.3- Expected Values for Continuous Random Variables

Definition 4.5

The expected value of a continuous random variable Y is:

$$E(Y) = \int_{-\infty}^{\infty} |y| f(y)\, dy \quad < \infty$$

red

provided that the integral exist.

Theorem 4.4

Let $g(Y)$ be a function of Y; Then the expected value of $g(Y)$ is given by:

$$E[g(Y)] = \int_{-\infty}^{\infty} g(y) f(y)\, dy.$$

provided that the integral exist.

THEOREM **4.5**

Let c be a constant and let $g(Y),\ g_1(Y),\ g_2(Y), \ldots, g_k(Y)$ be functions of a continuous random variable Y. Then the following results hold:

1. $E(c) = c$.
2. $E[cg(Y)] = cE[g(Y)]$.
3. $E[g_1(Y)+g_2(Y)+\cdots+g_k(Y)] = E[g_1(Y)]+E[g_2(Y)]+\cdots+E[g_k(Y)]$.

4.4 - The Uniform Probability Distribution

Definition 4.6

If $\theta_1 < \theta_2$, a random variable Y is said to have a continuous uniform probability distribution on the interval (θ_1, θ_2), if and only if the density function of Y is:

parameters.

$$f(y) = \begin{cases} \dfrac{1}{\theta_1 - \theta_2} & (\theta_1 < y < \theta_2) \\ 0 & \text{elsewhere.} \end{cases}$$

<u>Note</u>:

Computer random generator

↳ If we desire a set of observations on a random variable Y, with distribution function $F(y)$, we often can obtain the desired results by transforming a set of observations on a uniform random variable

Theorem 4.6

THEOREM 4.6 If $\theta_1 < \theta_2$ and Y is a random variable uniformly distributed on the interval (θ_1, θ_2), then

$$\mu = E(Y) = \frac{\theta_1 + \theta_2}{2} \quad \text{and} \quad \sigma^2 = V(Y) = \frac{(\theta_2 - \theta_1)^2}{12}.$$

Proof By Definition 4.5,

$$\begin{aligned} E(Y) &= \int_{-\infty}^{\infty} y f(y)\, dy \\ &= \int_{\theta_1}^{\theta_2} y \left(\frac{1}{\theta_2 - \theta_1} \right) dy \\ &= \left(\frac{1}{\theta_2 - \theta_1} \right) \frac{y^2}{2} \Bigg]_{\theta_1}^{\theta_2} = \frac{\theta_2^2 - \theta_1^2}{2(\theta_2 - \theta_1)} \\ &= \frac{\theta_2 + \theta_1}{2}. \end{aligned}$$

Note that the mean of a uniform random variable is simply the value midway between the two parameter values, θ_1 and θ_2. The derivation of the variance is left as an exercise.

4.5 - The Normal Probability Distribution

Definition 4.8

A random variable Y is said to have a normal probability distribution if and only if, $\sigma > 0$, & $-\infty < \mu < \infty$, the density function of Y is:

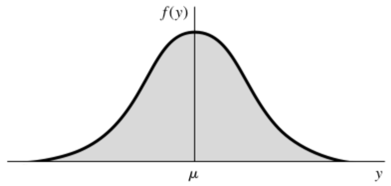
$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(y-\mu)^2/(2\sigma^2)} \qquad -\infty < y < \infty$$

<u>Theorem 4.7:</u>

If Y is a normally distributed random variable with parameters μ & σ, then

$$E(Y) = \mu \quad \& \quad V(Y) = \sigma^2$$

FIGURE 4.10 The normal probability density function



- We can always transform a normal random variable Y to a standard normal random variable Z by using:

$$Z = \frac{Y - \mu}{\sigma}$$

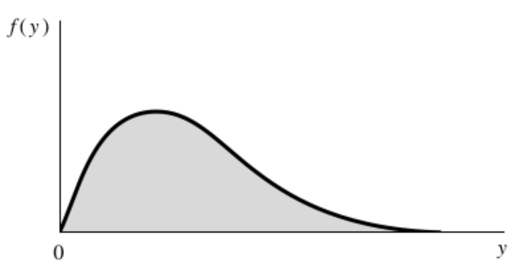
* Z locates a point measured from the mean of a normal random variable, with the distance expressed in units of standard deviation of the original normal random variable.

∴ The mean value Z must be 0, and its standard deviation must be equal to 1.

4.6- The Gamma Probability Distribution

- Some random variables are always non-negative and for various reasons yield distributions of data that are skewed (non-symmetric) to the right

FIGURE 4.15
A skewed probability density function



↳ * Most of the area under the density function is located near the origin.

* Density function drops gradually as y increases

Examples

- The lengths of time between malfunctions for aircraft engines.
- Time between arrivals at a supermarket checkout queue.
-

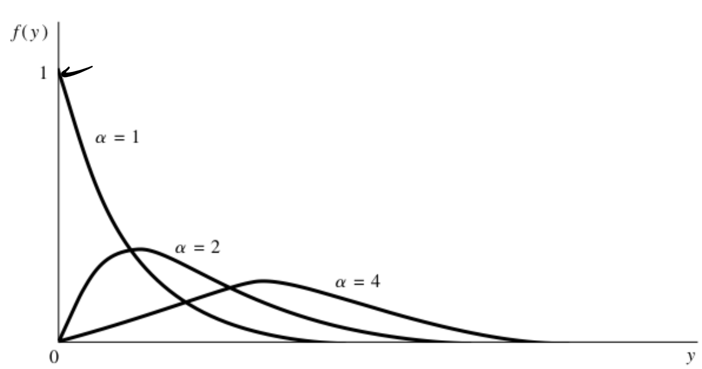
DEFINITION 4.9 A random variable Y is said to have a *gamma distribution with parameters* $\alpha > 0$ and $\beta > 0$ if and only if the density function of Y is

$$f(y) = \begin{cases} \dfrac{y^{\alpha-1}e^{-y/\beta}}{\beta^{\alpha}\Gamma(\alpha)}, & 0 \le y < \infty, \\ 0, & \text{elsewhere,} \end{cases}$$

where

$$\Gamma(\alpha) = \int_0^{\infty} y^{\alpha-1}e^{-y}\, dy.$$

FIGURE **4.16** Gamma density functions, $\beta = 1$



(More pg 211 - 219)

4-7 - The Beta Probability Distribution

- 2 parameter density function defined over the closed interval $0 \leq y \leq 1$.
- Used:
  - Model for proportions

eg: proportion of impurities in a chemical product

DEFINITION 4.12 A random variable Y is said to have a *beta probability distribution with parameters* $\alpha > 0$ *and* $\beta > 0$ if and only if the density function of Y is

$$f(y) = \begin{cases} \dfrac{y^{\alpha-1}(1-y)^{\beta-1}}{B(\alpha, \beta)}, & 0 \leq y \leq 1, \\ 0, & \text{elsewhere,} \end{cases}$$

where

$$B(\alpha, \beta) = \int_0^1 y^{\alpha-1}(1-y)^{\beta-1}\, dy = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$$

EXAMPLE 4.11 A gasoline wholesale distributor has bulk storage tanks that hold fixed supplies and are filled every Monday. Of interest to the wholesaler is the proportion of this supply that is sold during the week. Over many weeks of observation, the distributor found that this proportion could be modeled by a beta distribution with $\alpha = 4$ and $\beta = 2$. Find the probability that the wholesaler will sell at least 90% of her stock in a given week.

Solution If Y denotes the proportion sold during the week, then

$$f(y) = \begin{cases} \dfrac{\Gamma(4+2)}{\Gamma(4)\Gamma(2)} y^3(1-y), & 0 \le y \le 1, \\ 0, & \text{elsewhere,} \end{cases}$$

and

$$P(Y > .9) = \int_{.9}^{\infty} f(y)\,dy = \int_{.9}^{1} 20(y^3 - y^4)\,dy$$

$$= 20\left\{ \left.\frac{y^4}{4}\right]_{.9}^{1} - \left.\frac{y^5}{5}\right]_{.9}^{1} \right\} = 20(.004) = .08.$$

It is *not* very likely that 90% of the stock will be sold in a given week. ■

4.8 - General Comments

- Keep in mind that density functions are theoretical models for population of real data that occur in random phenomena.

* How do we we know which model?
* How much does it matter if we use the wrong density as our model for

reality?

↳ Unlikely ever to select a density function that provides a perfect representation.
[* goodness of fit is not the criterion for assessing the adequacy of our model].

* A good model is one that yields good inferences about the population of interest.

→ Selecting a reasonable model is sometimes a matter of acting on theoretical considerations.

eg Poisson random variable is indicated by the random behaviour of events in time.

→ $2^{\underline{nd}}$ selecting a model is to form a frequency histogram for data drawn from the population and to choose a density function that would usually appear to give a similar frequency curve.

Not all model selection is completely subjective. Statistical procedures are available to test a hypothesis that a population frequency distribution is of a particular type. We can also calculate a measure of goodness of fit for several distributions and select the best. Studies of many common inferential methods have been made to determine the magnitude of the errors of inference introduced by incorrect population models. It is comforting to know that many statistical methods of inference are insensitive to assumptions about the form of the underlying population frequency distribution.

The uniform, normal, gamma, and beta distributions offer an assortment of density functions that fit many population frequency distributions. Another, the Weibull distribution, appears in the exercises at the end of the chapter.

4.9 - Other Expected Values

4.10 - Tchebysheff's Theorem

- As was the case for discrete random variables, an interpretation of μ and σ for continuous random variables → provided by the empirical rule and Tchebysheff's theorem

THEOREM **4.13**

Tchebysheff's Theorem Let Y be a random variable with finite mean μ and variance σ^2. Then, for any $k > 0$,

$$P(|Y - \mu| < k\sigma) \geq 1 - \frac{1}{k^2} \quad \text{or} \quad P(|Y - \mu| \geq k\sigma) \leq \frac{1}{k^2}.$$

↳ Tchebysheff's Theorem enables us to find bounds for probabilities that ordinarily would have been obtained by tedious mathematical manipulations [integration or summation].

4.11 - Expectations of Discontinuous Functions and Mixed Probability Distributions.

- Some problems in probability & statistics sometimes involve functions that are partly continuous & partly discrete.

EXAMPLE 4.18 A retailer for a petroleum product sells a random amount Y each day. Suppose that Y, measured in thousands of gallons, has the probability density function

$$f(y) = \begin{cases} (3/8)y^2, & 0 \le y \le 2, \\ 0, & \text{elsewhere.} \end{cases}$$

The retailer's profit turns out to be \$100 for each 1000 gallons sold (10¢ per gallon) if $Y \le 1$ and \$40 *extra* per 1000 gallons (an extra 4¢ per gallon) if $Y > 1$. Find the retailer's expected profit for any given day.

Solution Let $g(Y)$ denote the retailer's daily profit. Then

$$g(Y) = \begin{cases} 100Y, & 0 \le Y \le 1, \\ 140Y, & 1 < Y \le 2. \end{cases}$$

We want to find expected profit; by Theorem 4.4, the expectation is

$$\begin{aligned} E[g(Y)] &= \int_{-\infty}^{\infty} g(y)f(y)\,dy \\ &= \int_0^1 100y\left[\left(\frac{3}{8}\right)y^2\right] dy + \int_1^2 140y\left[\left(\frac{3}{8}\right)y^2\right] dy \\ &= \frac{300}{(8)(4)}y^4\Big]_0^1 + \frac{420}{(8)(4)}y^4\Big]_1^2 \\ &= \frac{300}{32}(1) + \frac{420}{32}(15) = 206.25. \end{aligned}$$

Thus, the retailer can expect a profit of \$206.25 on the daily sale of this particular product. ■

An easy method for finding expectations of random variables with mixed distributions is given in Definition 4.15.

DEFINITION 4.15 Let Y have the mixed distribution function

$$F(y) = c_1F_1(y) + c_2F_2(y)$$

and suppose that X_1 is a discrete random variable with distribution function $F_1(y)$ and that X_2 is a continuous random variable with distribution function $F_2(y)$. Let $g(Y)$ denote a function of Y. Then

$$E[g(Y)] = c_1E[g(X_1)] + c_2E[g(X_2)].$$

4.12 - Summary

- This chapter presented probabilistic models for continuous random variables
- The density function, which provides a model for a population frequency distribution associated with a continuous random variable
↳ yields a mechanism for inferring characteristics of the population based on measurements contained in a sample taken from that population
↳ As a result, the density function provides a model for real distribution of data that exist or could be generated by repeated experimentation

The latter part of the chapter concerned expectations, particularly moments and moment-generating functions. It is important to focus attention on the reason for presenting these quantities and to avoid excessive concentration on the mathematical aspects of the material. Moments, particularly the mean and variance, are numerical descriptive measures for random variables. Particularly, we will subsequently see that it is sometimes difficult to find the probability distribution for a random variable Y or a function $g(Y)$, and we already have observed that integration over intervals for many density functions (the normal and gamma, for example) is very difficult. When this occurs, we can approximately describe the behavior of the random variable by using its moments along with Tchebysheff's theorem and the empirical rule (Chapter 1).